JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **160**, 314–322 (1991)

A Duality Approach to Best Uniform Convex Approximation

S. E. WEINSTEIN AND YUESHENG XU*

Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia 23529

Submitted by V. Lakshmikantham

Received June 27, 1989

1. INTRODUCTION

Let $C[a, b]$ be the space of continuous functions on $[a, b]$ endowed with the uniform norm $\|f\|_\infty = \sup\{|f(x)| : x \in [a, b]\}$. Let K be the set of convex functions defined on $[a, b]$. A function $g^* \in K$ is said to be a best uniform convex approximation to $f \in C[a, b]$ if

$$\|f - g^*\|_\infty = \inf\{\|f - g\|_\infty : g \in K\}. \tag{1.1}$$

The existence of a best uniform convex approximation to a bounded function was demonstrated in [3], where an algorithm for the computation of a best approximation by means of linear programming was also presented. The characterization of alternant-type is a special case of a result announced in [1] and proved in [8]. In this paper, the term "best approximation" means best uniform convex approximation unless stated otherwise. We establish a duality theorem that expresses the error of the best approximation in terms of the supremum of a linear functional of f and use this duality to investigate the properties of best approximations. We use this duality result to obtain bounds for the error of best approximation, to give an alternative proof to the characterization of the best approximation, and to characterize the set of linear negative alternants. We also define a "functional interval" (similar to that defined in [4] for monotone approximation) which we show is a necessary condition for best convex approximation.

A similar duality approach has been used in [4, 7] to investigate best monotone approximation and best quasiconvex approximation, respectively.

* Current address: Department of Mathematics, North Dakota State University, Fargo, ND 58105.

0022-247X/91 $3.00

2. Duality

Define

$$S = \{(x, y; \lambda) : x, y \in [a, b], 0 \leqslant \lambda \leqslant 1\}. \tag{2.1}$$

S is a compact set in R^3. For $f \in C[a, b]$, define the function F on S by

$$F(x, y, \lambda) = (-1/2)[\lambda f(x) - f(\lambda x + (1-\lambda) y) + (1-\lambda) f(y)]. \tag{2.2}$$

Let

$$\delta = \delta(f) = \sup\{F(x, y, \lambda) : (x, y; \lambda) \in S\}. \tag{2.3}$$

δ is a measure of the convexity of the function f. We see in Lemma 1 that $\delta = 0$ is equivalent to f being convex. Let

$$\Delta = \{(x, y; \lambda) \in S : F(x, y, \lambda) = \delta\}. \tag{2.4}$$

Since f is continuous on $[a, b]$, F is continuous on S. Thus, F assumes its maximum on S, and therefore Δ is nonempty. For $f \in C[a, b]$, define *the greatest convex minorant* or *lower convex envelope* of f by

$$\operatorname{env} f(t) = \sup\{g(t) : g \in K \text{ and } f \geqslant g \text{ on } [a, b]\}, \qquad t \in [a, b], \tag{2.5}$$

where $f \geqslant g$ on $[a, b]$ means that $f(s) \geqslant g(s)$ for all $s \in [a, b]$. We remark that $\operatorname{env} f$ is the largest continuous convex function that does not exceed f at any point in $[a, b]$ (see [2]).

Lemma 1. *Let $f \in C[a, b]$. Then, $\delta = 0$ if and only if f is convex.*

Proof. If f is convex, then for all $(x, y; \lambda) \in S$, $F(x, y, \lambda) \leqslant 0$. Hence, $\delta = 0$. Conversely, if f is not convex, then there exists $(x, y; \lambda) \in S$ with $x \neq y$ and $0 < \lambda < 1$ such that $F(x, y, \lambda) > 0$. Thus, $\delta > 0$.

Lemma 2. *Let $f \in C[a, b] - K$. If $(x, y; \lambda) \in \Delta$, then $x \neq y$ and $0 < \lambda < 1$.*

Proof. Assume to the contrary that one of the following statements is true: $x = y$, $\lambda = 0$, or $\lambda = 1$. Thus $F(x, y, \lambda) = 0$. Since $(x, y; \lambda) \in \Delta$, $\delta = F(x, y, \lambda) = 0$. By Lemma 1, f is convex. This contradicts the hypothesis.

The following lemma was basically proved in [3]:

Lemma 3. *Let $f \in C[a, b]$. Then, $f(a) = \operatorname{env} f(a)$ and $f(b) = \operatorname{env} f(b)$.*

Now we can establish a duality theorem showing that $\delta(f)$ is the error of best approximation.

THEOREM 1 (Duality). *Let $f \in C[a, b]$. Then,*

$$\inf\{\|f-g\|_\infty : g \in K\} = \delta(f). \tag{2.6}$$

Proof. For any $(x, y; \lambda) \in S$ and all $g \in K$,

$$\lambda g(x) - g(\lambda x + (1-\lambda) y) + (1-\lambda) g(y) \geqslant 0,$$

and thus

$$F(x, y, \lambda) \leqslant F(x, y, \lambda) + (1/2)[\lambda g(x) - g(\lambda x + (1-\lambda) y) + (1-\lambda) g(y)] \leqslant \|f-g\|_\infty.$$

Consequently, $\delta(f) \leqslant \inf\{\|f-g\|_\infty : g \in K\}$.

To complete this proof, let

$$\bar{g}(t) = \operatorname{env} f(t) + \delta(f), \qquad \text{for all} \quad t \in [a, b]. \tag{2.7}$$

Since $\operatorname{env} f \leqslant f$, on $[a, b]$, we have $\bar{g}(t) \leqslant f(t) + \delta(f)$, for all $t \in [a, b]$. Assume that there exists an $x_0 \in (a, b)$ such that

$$f(x_0) - \delta(f) > \bar{g}(x_0) = \operatorname{env} f(x_0) + \delta(f). \tag{2.8}$$

By virtue of the continuity of $f - \operatorname{env} f$, there exists some open interval $I \subset [a, b]$ such that

$$f(t) - \operatorname{env} f(t) > 2\delta(f), \qquad \text{for all} \quad t \in I.$$

Lemma 3 then implies that there exist $x_1, x_2 \in [a, b]$ with $I \subseteq (x_1, x_2)$ such that

$$f(x_1) = \operatorname{env} f(x_1), \qquad f(x_2) = \operatorname{env} f(x_2),$$

and

$$f(t) - \operatorname{env} f(t) > 0 \qquad \text{for all} \quad t \in (x_1, x_2).$$

By a similar reasoning as in [3], we can show that $\operatorname{env} f$ is linear on (x_1, x_2). Therefore, for some $\lambda_0 \in (0, 1)$, $x_0 = \lambda_0 x_1 + (1-\lambda_0) x_2$ and

$$\operatorname{env} f(\lambda_0 x_1 + (1-\lambda_0) x_2) = \lambda_0 f(x_1) + (1-\lambda_0) f(x_2).$$

It follows that

$$\begin{aligned} &f(\lambda_0 x_1 + (1-\lambda_0) x_2) - \operatorname{env} f(\lambda_0 x_1 + (1-\lambda_0) x_2) \\ &\quad = \lambda_0 f(x_1) + f(\lambda_0 x_1 + (1-\lambda_0) x_2) - (1-\lambda_0) f(x_2) \\ &\quad > 2\delta(f). \end{aligned}$$

This last inequality contradicts the definition of $\delta(f)$. This contradiction implies that (2.8) cannot hold. Thus

$$\bar{g}(t) \geqslant f(t) - \delta(f), \qquad \text{for all} \quad t \in [a, b].$$

Hence,

$$f(t) - \delta(f) \leqslant \bar{g}(t) \leqslant f(t) + \delta(f), \qquad \text{for all} \quad t \in [a, b].$$

Since $\bar{g} \in K$, we have established Eq. (2.6).

COROLLARY 1. *Let $f \in C[a, b]$. Then, $\bar{g} = \operatorname{env} f + \delta(f)$ is a best convex approximation to f*

THEOREM 2. *Let $\rho(f) = \inf\{\|f - g\|_\infty : g \in K\}$.*

(i) *Let $f \in C^1[a, b]$. Then,*

$$\rho(f) \leqslant [(b-a)/8] \sup\{f'(x) - f'(y) : a \leqslant x \leqslant y \leqslant b\}. \tag{2.9}$$

(ii) *Let $f \in C^2[a, b]$. Then,*

$$\rho(f) \leqslant [(b-a)^2/16] \sup\{[-f''(x)]_+ : x \in [a, b]\}, \tag{2.10}$$

where

$$[a]_+ = \begin{cases} 0 & \textit{if} \quad a \leqslant 0 \\ a & \textit{if} \quad a > 0. \end{cases}$$

Proof. (i) Note that for $(x, y; \lambda) \in S$,

$$\begin{aligned} F(x, y, \lambda) = (1/2)\{&\lambda[f(\lambda x + (1-\lambda) y) - f(x)] \\ &+ (1-\lambda)[f(\lambda x + (1-\lambda) y) - f(y)]\}. \end{aligned}$$

Since $f \in C^1[a, b]$,

$$F(x, y, \lambda) = (1/2)\, \lambda(1-\lambda)(y-x)[f'(t_1) - f'(t_2)],$$

for some $t_1 \in [x, \lambda x + (1-\lambda) y]$ and $t_2 \in [\lambda x + (1-\lambda) y, y]$. Therefore,

$$\delta(f) \leqslant [(b-a)/8] \sup\{f'(x) - f'(y) : a \leqslant x \leqslant y \leqslant b\}.$$

(ii) If $f \in C^2[a, b]$ and $(x, y; \lambda) \in S$, for some $t \in [x, y]$,

$$\begin{aligned} F(x, y, \lambda) &= (-1/2)\, \lambda(1-\lambda)(y-x)^2\, [x, \lambda x + (1-\lambda) y, y] f \\ &= (-1/4)\, \lambda(1-\lambda)(y-x)^2 f''(t), \end{aligned}$$

where $[t_1, t_2, t_3]f$ denotes the second divided difference of f at t_1, t_2, t_3. Hence, inequality (2.10) follows.

As another application of Theorem 1, we provide an alternative proof of the characterization of best convex approximation to a continuous function, which was announced in [1] and proved in [8].

CHARACTERIZATION THEOREM. *Let $f \in C[a,b]-K$. $g^* \in K$ is a best convex approximation to f if and only if there exist $x<y$ in $[a,b]$ and $\lambda \in (0,1)$ such that g^* is linear on $[x,y]$ and satisfies*

$$f(x)-g^*(x)=f(y)-g^*(y)=-\|f-g^*\|_\infty,$$

and

$$f(\lambda x+(1-\lambda)y)-g^*(\lambda x+(1-\lambda)y)=\|f-g^*\|_\infty.$$

Proof. (*Necessity*) By the hypothesis and Theorem 1, $\|f-g^*\|_\infty = \delta(f)$. In view of the continuity of f, Δ is nonempty. Assume $(x, y; \lambda) \in \Delta$. Then by Lemma 2, $x<y$ and $0<\lambda<1$. Since $g^* \in K$, the following inequality holds:

$$G(x,y,\lambda) \equiv (1/2)[\lambda g^*(x) - g^*(\lambda x+(1-\lambda)y)+(1-\lambda)g^*(y)] \geqslant 0.$$

If $G(x,y,\lambda)>0$, then

$$\begin{aligned}\delta(f) &= F(x,y,\lambda)\\ &< F(x,y,\lambda)+G(x,y,\lambda)\\ &\leqslant (1/2)[\lambda\,\|f-g^*\|_\infty + \|f-g^*\|_\infty + (1-\lambda)\,\|f-g^*\|_\infty]\\ &= \|f-g^*\|_\infty.\end{aligned}$$

This contradicts Theorem 1. Thus $G(x,y,\lambda)=0$. It follows from this equation and the convexity of g^* that g^* is linear on $[x,y]$. Therefore,

$$\begin{aligned}\delta(f) &= F(x,y,\lambda)+G(x,y,\lambda)\\ &= (1/2)\{\lambda[g^*(x)-f(x)] + [f(\lambda x+(1-\lambda)y)-g^*(\lambda x+(1-\lambda)y)]\\ &\quad + (1-\lambda)[g^*(y)-f(y)]\}.\end{aligned}$$

However, from Theorem 1, we have

$$-\delta(f) \leqslant f(t)-g^*(t) \leqslant \delta(f), \qquad \text{for all} \quad t \in [a,b].$$

If $g^*(x) - f(x) < \delta(f)$, then

$$\begin{aligned}(1-\lambda/2)\,\delta(f) &< (1/2)\{[f(\lambda x+(1-\lambda)\,y)\\ &\qquad - g^*(\lambda x+(1-\lambda)\,y)] + (1-\lambda)[g^*(y)-f(y)]\}\\ &\leqslant (1-\lambda/2)\,\|f-g^*\|_\infty,\end{aligned}$$

and thus $\delta(f) < \|f-g^*\|_\infty$, which is a contradiction. This contradiction implies that $g^*(x) - f(x) = \delta(f)$. Similarly, we can show that $f(\lambda x + (1-\lambda)\,y) - g^*(\lambda x + (1-\lambda)\,y) = \delta(f)$ and $g^*(y) - f((y) = \delta(f)$. These three equations and Theorem 1 establish the necessity of the characterization.

(*Sufficiency*) From the assumptions we have

$$\begin{aligned}\delta(f) &\geqslant F(x, y, \lambda)\\ &= \|f-g^*\|_\infty + (1/2)[-\lambda g^*(x)\\ &\qquad + g^*(\lambda x+(1-\lambda)\,y) - (1-\lambda)\,g^*(y)]\\ &= \|f-g^*\|_\infty,\end{aligned}$$

where the last equality holds because of the linearity of g^* on $[x, y]$. Hence, by Theorem 1, g^* is a best convex approximation to f on $[a, b]$.

3. Some Properties of Best Convex Approximations

In this section, we characterize the set of linear negative alternants of $f - g^*$, where g^* is a best convex approximation to $f \in C[a, b]$ and identify two functions which are respectively a lower bound and an upper bound of any best approximation to f.

For a real-valued function h defined on $[a, b]$, $a \leqslant x_1 < x_2 < x_3 \leqslant b$ is said to be a *negative alternant* of h, if $-h(x_1) = h(x_2) = -h(x_3) = \|h\|_\infty$. For $f \in C[a, b] - K$ and $g \in K$, define the set of *linear negative alternants* of $f - g$ by

$$\begin{aligned}A(f-g) = \{&(x, y; \lambda) \in S : g \text{ is linear on } [x, y] \text{ and}\\ &x < \lambda x + (1-\lambda)\,y < y \text{ is a negative alternant of } f-g\}.\end{aligned}$$

The following theorem characterizes the set of linear negative alternants of $f - g^*$, where g^* is a best convex approximation to f.

THEOREM 3. *Let $f \in C[a, b] - K$ and let g^* be a best convex approximation to f on $[a, b]$. Then,*

$$A(f - g^*) = \Delta.$$

Proof. Let $(x, y; \lambda) \in \Delta$. By a similar reasoning as in the proof of the characterization of best convex approximation, we find $(x, y; \lambda) \in A(f - g^*)$. This gives $\Delta \subseteq A(f - g^*)$. Conversely, assume $(x, y; \lambda) \in A(f - g^*)$. Then, g^* is linear on $[x, y]$ and satisfies

$$f(x) - g^*(x) = f(y) - g^*(y) = -\|f - g^*\|_\infty = -\delta(f),$$

and

$$f(\lambda x + (1 - \lambda) y) - g^*(\lambda x + (1 - \lambda) y) = \|f - g^*\|_\infty = \delta(f).$$

Hence,

$$\begin{aligned} \delta(f) &\geq F(x, y, \lambda) \\ &= (1/2)\{-\lambda[g^*(x) - \delta(f)] + [g^*(\lambda x + (1 - \lambda) y) - \delta(f)] \\ &\quad - (1 - \lambda)[g^*(y) - \delta(f)]\} = \delta(f). \end{aligned}$$

This implies that $F(x, y, \lambda) = \delta(f)$, and thus $(x, y; \lambda) \in \Delta$. Accordingly, $A(f - g^*) = \Delta$.

COROLLARY 2. *Let $f \in C[a, b] - K$ and let g^* be a best convex approximation to f on $[a, b]$. Then, for all $(x, y; \lambda) \in \Delta$ with $x < y$, g^* is linear on $[x, y]$ and*

$$g^*(\mu x + (1 - \mu) y) = \mu f(x) + (1 - \mu) f(y) + \delta(f), \qquad \mu \in [0, 1].$$

Proof. For $(x, y; \lambda) \in \Delta$, by Theorem 3, $(x, y; \lambda) \in A(f - g^*)$. Hence, g^* is linear on $[x, y]$, and $g^*(x) = f(x) + \delta(f)$ and $g^*(y) = f(y) + \delta(f)$. Therefore, by linear interpolation, for all $\mu \in [0, 1]$,

$$\begin{aligned} &g^*(\mu x + (1 - \mu) y) \\ &\quad = g^*(x)[\mu x + (1 - \mu) y - y]/(x - y) \\ &\qquad + g^*(y)[\mu x + (1 - \mu) y - x]/(y - x) \\ &\quad = f(x) \mu + f(y)(1 - \mu) + \delta(f). \end{aligned}$$

Now we identify two functions which bound any best approximation.

For $(x, y; \lambda) \in \Delta$, denote the linear interpolant to $(x, f(x)+\delta(f))$ and $(y, f(y)+\delta(f))$ on $[a, b]$ by

$$l[x, y](t) = f(x)(t-y)/(x-y) + f(y)(t-x)/(y-x) + \delta(f), \qquad t \in [a, b].$$

Let

$$L = \{l[x, y] : (x, y; \lambda) \in \Delta\}, \tag{3.1}$$

and

$$G = \{\operatorname{env} f - \delta(f)\} \cup L. \tag{3.2}$$

Define

$$\underline{g}(t) = \sup\{g(t) : g \in G\}, \qquad t \in [a, b]. \tag{3.3}$$

It is easy to verify that $\underline{g}$ is a convex function on $[a, b]$ and if f is convex then $\underline{g} = f$. The next theorem shows that this convex function is a lower bound of the best approximations to f.

THEOREM 4. *Let $f \in C[a, b]$. If $g^* \in K$ is a best convex approximation to f, then,*

$$\underline{g}(t) \leqslant g^*(t) \leqslant \bar{g}(t), \qquad \textit{for all} \quad t \in [a, b], \tag{3.4}$$

where $\bar{g}$ was defined in (2.7).

Proof. In [3], it has been proved that $g^*(t) \leqslant \operatorname{env} f(t) + \|f - g^*\|_\infty$. By replacing $\|f - g^*\|_\infty$ by $\delta(f)$, we obtain the upper bound. To show the lower bound, assume to the contrary that there exists some $z \in [a, b]$ such that $\underline{g}(z) > g^*(z)$. Define

$$P = \bigcup \{[x, y] : (x, y; \lambda) \in \Delta\}. \tag{3.5}$$

By the definition of $\underline{g}$,

$$\underline{g}(t) = g^*(t), \qquad \text{for all} \quad t \in P.$$

Hence z is not in P. If $\underline{g}(z) = \operatorname{env} f(z) - \delta(f)$, then

$$f((z) - \delta(f) \geqslant \operatorname{env} f(z) - \delta(f) > g^*(z),$$

which contradicts the hypothesis that g^* is a best convex approximation to f. Therefore, there exists $(x, y; \lambda) \in \Delta$, such that $l[x, y](z) > g^*(z)$. This contradicts the convexity of g^*. It follows that $\underline{g}(t) \leqslant g^*(t)$, for all $t \in [a, b]$.

As shown in Corollary 1, $\bar{g}$ is a best convex approximation to f. Hence it is the greatest best convex approximation to f. However, $\underline{g}$ may not be a best convex approximation to f.

It is shown in [4, 6] that if f is continuous but not nondecreasing on $[a, b]$, then there exists a best monotone approximation to f that is in C^{∞}. However, an analogous statement is not true for best convex approximation. To see this, let us consider the following example: Assume

$$f(x)=\begin{cases} -6x+1 & 0\leqslant x\leqslant 1/8 \\ 4x-1/4 & 1/8<x\leqslant 1/4 \\ -3x+3/2 & 1/4<x\leqslant 1/2 \\ 3x-3/2 & 1/2<x\leqslant 3/4 \\ -4x+15/4 & 3/4<x\leqslant 7/8 \\ 6x-5 & 7/8<x\leqslant 1. \end{cases}$$

Then f is continuous but is not convex on $[0, 1]$. $\delta(f)=\frac{5}{16}$ and

$$\Delta=\{(1/8, 1/2; 1/2), (1/2, 7/8; 1/2)\}.$$

Hence, every best convex approximation has a knot at $x=\frac{1}{2}$, and thus is not differentiable at $\frac{1}{2}$.

References

1. H. G. Burchard, Extremal positive splines with applications to interpolation and approximation by generalized convex functions, *Bull. Amer. Math. Soc.* **79** (1973), 959–963.
2. A. W. Roberts and D. E. Varberg, "Convex Function," Academic Press, New York, 1973.
3. V. A. Ubhaya, An $O(n)$ algorithm for discrete n-point convex approximation with applications to continuous case, *J. Math. Anal. Appl.* **72** (1979), 338–354.
4. V. A. Ubhaya, Isotone optimization. I, *J. Approx. Theory* **12** (1974), 146–159.
5. V. A. Ubhaya, Isotone optimization. II, *J. Approx. Theory* **12** (1974), 315–331.
6. V. A. Ubhaya, Moduli of monotonicity with applications to monotone polynomial approximation, *SIAM J. Math. Anal.* **7** (1976), 117–130.
7. S. E. Weinstein and Yuesheng Xu, Best quasi-convex uniform approximation, *J. Math. Anal. Appl.* **152** (1990), 240–251.
8. D. Zwick, Best approximation by convex functions, *Amer. Math. Monthly* **94** (1987), 528–534.